PRELIMINARY ACOUSTIC ANALYSIS OF NOISE COMPONENTS IN PATIENTS WITH PARKINSON'S DISEASE

Zoltán Galáž
Doctoral Degree Programme (1), FEEC BUT
E-mail: z.galaz@phd.feec.vutbr.cz

Supervised by: Petr Sysel
E-mail: sysel@feec.vutbr.cz

Abstract: This paper deals with acoustic analysis of noise components extracted from speech signals of patients with Parkinson's disease (PD) who recited a poem. Experimental dataset consisted of 97 PD patients with different disease progress and 55 healthy controls (HC). The analysis is based on parametrization of 2 rhymes recitation using dysphonia features. We obtained classification accuracy 76.66 % for female speakers, 69.65 % for male speakers and 69.24 % for the mixture of both genders.

Keywords: Parkinson's disease, Empirical Mode Decomposition, hypokinetic dysarthria, dysphonia

1 INTRODUCTION

Parkinson's disease (PD) is the world's second most frequent neurodegenerative disease [1]. PD was firstly described in 1817 by James Parkinson in his essay "An Essay on the Shaking Palsy". It is a chronic disorder that is characterized by progressive loss of dopaminergic neurons in part of the brain called *substancia nigra pars compacta* [2]. Deficiency of dopamine is considered to be major cause of parkinsonian symptoms [3]. In addition to motor symptoms such as resting tremor, bradykinesia, muscular rigidity and postural instability, patients with PD also develop behavioural alternations, reduction of cognitive abilities and the world-wide studies recently showed that one of the most frequent manifestation of PD is speech disorder. Speech deficits associated with PD are described as dysphonia (inability to produce normal vocal sounds including hoarseness and roughness, which reduce the intelligibility of speech) and dysarthria (difficulty in pronouncing words), especially its form called hypokinetic dysarthria (HD) [4].

This paper is focused mainly on the analysis of voice quality degradation caused by the fact that due to vocal chord's muscle stiffness in patients with PD, the turbulent flow from lungs is characterized by significant fluctuations, which is reflected in voice tremor [5], and also in presence of higher amount of noise caused by incomplete vocal fold closures. Features that quantify the speech impairments are generally called dysphonia measures, see [6]. An initial assumption in this paper is that due to speech impairments in patients with PD, features that quantify the noise components in the speech signal will provide better clinical information about the state of the vocal degradation. Therefore we assume that these features are good candidates for PD and healthy speech discrimination.

The rest of this paper is organized as follows. Section 2 presents the dataset and the methods used in this study, description of features used to quantify noise components and statistical analysis of these features. Experimental results are presented in section 3, and finally section 4 provides discussion and some conclusions.

2 MATERIALS AND METHODS

2.1 Patients and Data Acquisition

A grand total of 152 Czech native speakers were studied. Altogether, 97 PD patients (53 men/44 women; mean age 67.52 ± 8.29 years; mean disease duration 7.80 ± 4.42 years; UDPRS III score[1] 24.91 ± 11.97; LED[2] 1005.93 ± 545.66 mg, and 55 healthy controls (29 men/22 women; mean age 63.96 ± 9.21 years) were enrolled at the First Department of Neurology, St. Anne's University Hospital in Brno, Czech Republic. The healthy participants had no history or presence of speech disorders or brain diseases, including neurological and psychiatric illnesses. Every speaker was asked to recite a poem that consists of 2 rhymes:

in Czech	English translation
Chcete vidět velký lov?	*Would you like to see a big hunt?*
Budu lovit v džungli slov.	*I will be hunting in a jungle of words.*
Osedlám si Pegasa,	*I will saddle Pegasus,*
chýtím báseň do lasa!	*I will catch the poem into a lasso.*

All speakers read the poem on a paper and tried to recite it for themselves. Afterwards, they recited the poem into a microphone. Speech signals were sampled with sampling frequency $f_s = 48\,\text{kHz}$ and consequently downsampled to 16 kHz. All participants signed an informed consent form that had been approved by the Ethics Committee of St. Anne's University Hospital in Brno.

2.2 Speech features extraction

In this paper we used signal-to-noise ratio derived from discrete time wavelet transform (SNR DTW), harmonics-to-noise ratio (HNR), noise-to-harmonic ratio (NHR), normalized noise energy (NNE), modulation energy (ME), glottal-to-noise excitation ratio (GNE) and detrended fluctuation analysis (DFA). These dysphonia measures are often used to evaluate the noise components in speech signals and they have been used several times for detection of laryngeal pathologies. We also used features based on the theory of empirical mode decomposition (EMD) to decompose the speech signal into small number of intrinsic mode functions (IMF). Based on consideration that time-varying high frequency components represented by the first few IMFs are related to instability of vocal folds vibration, Tsanas et al. proposed several SNR (signal-to-noise ratio) and NSR (noise-to-signal ratio) measures [6]. Recently we proposed several features based on IMFs, e.g. glottal-to-noise excitation ratio extracted from the first IMF (IMF-GNE). For the purpose of speech feature extraction, we used Praat acoustic analysis software [7] and Neurological Disorders Analysis Tool (NDAT) written in MATLAB and developed at the Brno University of Technology.

Finally, to obtain complete statistical representation of available features, we computed several statistical functionals of feature vectors. We included descriptive statistics (max, min, position of max, position of min, relative pos. of max, relative pos. of min), range characteristics (range, relative range, interquartile range, rel. interquartile range, interdecile range, rel. interdecile range, interpercentile range, rel. interpercentile range, studentized range), also several means (arithmetic mean, geometric mean, trimmed means without outliers (5 %, 10 %, 20 %, 30 %, 40 %, 50 %)), quartiles (25/lower, 75/upper), percentiles (1st, 5th, 10th, 20th, 30th, 40th, 60th, 70th, 80th, 90th, 95th and 99th), moments (3rd, 4th, 5th and 6th, kurtosis, skewness, Spearman's 1st and 2nd skewness coefficient) and other (median, mode, var, std, mean absolute deviation, median absolute dev., geometric standard dev., coefficient of variation, index of dispersion, modulation, Shannon entropy, second-order Rényi entropy, slope, offset and error of linear regression) statistical functions.

[1] Unified Parkinson's disease rating scale, part III: evaluation of motor function

[2] L-dopa equivalent daily dose

2.3 Speech features analysis

A general problem in data analysis is so called *curse of dimensionality*: presence of very large number of features often inhibits detection of useful patterns underlying the data. In order to reduce the dimensionality of input feature set and remove non-relevant features, we used minimum Redundancy Maximum Relevance (mRMR) algorithm. We used implementation of Tsanas et al. [8] and computed mRMR using Spearman's correlation coefficient as a criterion to quantify statistical relationships between features and response (speakers' label). Subsequently we performed statistical analysis of our reduced feature space. We firstly calculated Spearman's rank sum correlation coefficients and normalised mutual information between feature vectors and associated diagnosis (HC vs. PD). Spearman's correlation coefficient (ρ) is a statistical measure of strength of a monotonic relationship between feature vectors and associated response. Mutual information (MI) is a measure of the amount of information shared by two random variables. The larger the value of MI is the stronger statistical association between the feature and the response can be observed.

Next, features were used separately as an input to Support vector machines (SVM) and Random Forests (RF) classifiers to evaluate their individual discrimination power. Resulting classification accuracies (mean of individual accuracies) are listed in Table 1. Finally, we performed Mann-Whitney U test to compare calculated features between healthy controls and PD patients. The Mann-Whitney U test is non-parametric statistical test assessing the difference (significant/insignificant) of two independent groups. Table 1 summarizes statistical properties of top five features with largest relevance to response sorted according to significance level calculated by the Mann-Whitney U test. We computed these statistics for each gender separately and also for a mixture of both. Mann-Whitney U test indicated significant differences ($p < 0.01$) between healthy controls and PD patients for all features listed in the table. Probability density estimations computed for the best three features are represented in Figure 1.

Features	p	ρ	MI	ACC [%]
		Female participants		
$\mathrm{CPP(1.\,IMF)}_{20^{th}\mathrm{percentile}}$	0.000020	0.1290	0.9727	55.2239
$\mathrm{FD(1.\,IMF)}_{\mathrm{mean\ -\ 10\,\%\ outliers}}$	**0.000028**	**-0.4729**	**0.8606**	**76.6567**
$\mathrm{FD(1.\,IMF)}_{\mathrm{harmonic\ mean}}$	0.000036	-0.1896	0.9017	52.2388
$\mathrm{TKEO(IMF\,NSR)}_{\mathrm{IQR}}$	0.000052	-0.3698	0.9727	55.2239
$\mathrm{TKEO(IMF\,NSR)}_{\mathrm{relative\ IQR}}$	0.000085	-0.4642	0.9727	62.6866
		Male participants		
$\mathrm{TKEO(IMF\,NSR)}_{\mathrm{kurtois}}$	**0.000473**	**0.4009**	**0.8884**	**69.6471**
$\mathrm{TKEO(IMF\,SNR)}_{60^{th}\mathrm{percentile}}$	0.001469	-0.2862	0.8884	54.7059
$\mathrm{SE(IMF\,SNR)}_{\mathrm{mean\ -\ 40\,\%\ outliers}}$	0.002168	-0.2955	0.8884	52.3529
$\mathrm{FD(1.\,IMF)}_{70^{th}\mathrm{percentile}}$	0.002238	-0.3195	0.6793	57.6471
$\mathrm{RE(IMF\,SNR)}_{\mathrm{mean\ -\ 10\,\%\ outliers}}$	0.002540	-0.2643	0.8884	55.8824
		All participants		
$\mathrm{TKEO(IMF\,SNR)}_{\mathrm{kurtois}}$	**0.000001**	**0.2295**	**0.9329**	**67.2368**
$\mathrm{TKEO(IMF\,SNR)}_{\mathrm{IQR}}$	0.000026	-0.2742	0.9329	53.6184
$\mathrm{TKEO(IMF\,SNR)}_{1^{st}\mathrm{percentile}}$	0.000028	-0.1084	0.9329	51.9737
$\mathrm{CPP(1.\,IMF)}_{20^{th}\mathrm{percentile}}$	0.000032	-0.0540	0.9329	53.2895
$\mathrm{TKEO(IMF\,NSR)}_{\mathrm{geometric\ std}}$	0.000035	0.2986	-0.3352	57.2368

Table 1: Statistical analysis of top five features selected by Mann-Whitney U-test. Table notation: IQR – interquantile range; FD – fractal dimension; SE – Shannon entropy; RE – Rényi entropy.

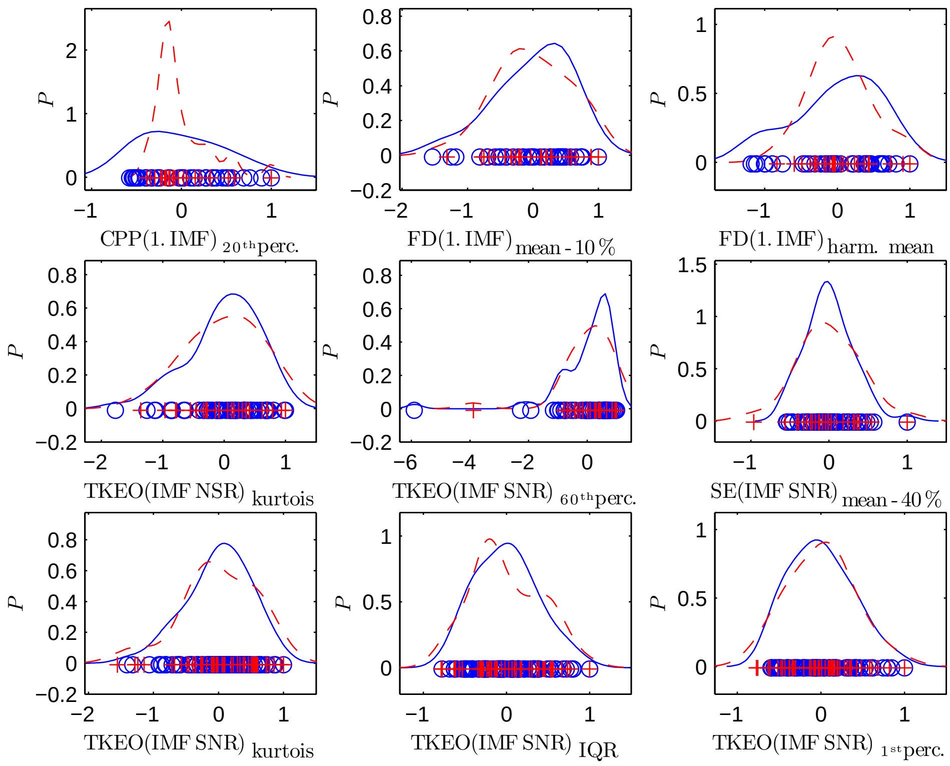

Figure 1: Density estimation plots (computed using density estimation with Gaussian kernels) of top three features selected by Mann-Whitney U test for both genders separately and for a mixture of both (First row of figures represents density estimations for female participants, the second row represents density estimations for male participants and the third one is related to a mixture of both genders). Red color represents healthy speech and blue color represents pathological one.

3 RESULTS

The aim of this study was to analyse the noise components present in the speech signal of patients with PD. We studied a poem recitation task of 97 PD patients and 55 gender and age matched controls. Some demographic and clinical data are summarized in Section 2. We found that patients with PD exhibit significant deficits in voice quality measures based on the theory of Empirical Mode Decomposition which indicates the presence of voice tremor due to vocal fold dysfunctions in combination with respiratory system impairments (breathy voice). Altogether, we calculated more than 1400 dysphonia features per subject.

Next, we computed the relevance of speech features to the diagnosis (HC/PD). The analysis of prosodic features by Spearman's rank correlation, mutual information, Mann-Whitney U test and single feature classification can be seen in Table 1. All selected features significantly correlated with the participants' label ($p < 0.001$). Figure 1 presents probability density estimation plots, computed using kernel density estimation with Gaussian kernels, of top three dysphonia measures. According to the computed classification accuracies given in Table 1 it is clear that female speakers were classified with higher accuracy than the male speakers or the combination of both genders. Based on the results of this study we conclude that underlying processes of degradation in PD speech may be different in male and female subjects.

4 CONCLUSION

In this paper we have analysed a poem recitation of patients with PD and healthy controls in order to estimate vocal impairments, which is a novel approach in the field of non-invasive PD diagnosis. We performed our measurements in male and female subjects separately and also in the mixture of both genders. It was shown that dysphonia features quantifying the noise components of speech signals can be used for diagnosis of PD with classification accuracy over 74 % using only one feature. Note that calculated classification accuracies represent only the discrimination power of a single feature, therefore any possible combination of extracted features is assumed to perform better. In addition, we have shown that splitting the data into male and female data subsets (data partitioning) reveals distinct PD speech characteristics in males and females and this tentatively suggests different pathological patterns in these two groups.

In our future study we will focus on deeper statistical analysis of noise components using dysphonia features. We will also use the sequential forward feature selection algorithm (SFFS) to select the best possible feature subset and to find the combination of speech features that provides maximal clinical information. We assume that the poem recitation task in combination with the analysis of noise components can be used not only for the binary classification purposes, but it can be also used for the purpose of disease progress tracking.

ACKNOWLEDGEMENT

Research described in this paper was financed by the National Sustainability Program under grant LO1401 and by projects NT13499 (Speech, its impairment and cognitive performance in Parkinson's disease), COST IC1206, project "CEITEC, Central European Institute of Technology": (CZ.1.05/1.1.00/02.0068) from the European Regional Development Fund. For the research, infrastructure of the SIX Center was used.

REFERENCES

[1] M. C. de Rijk et al., Prevalence of parkinson's disease in europe: A collaborative study of population-based cohorts, Neurology 54 (2000) 21–23.

[2] O. Hornykiewicz, Biochemical Aspects of Parkinson's Disease, Neurology 51 (1998) S2–S9.

[3] P. Brodal, The Central Nervous System: Structure and Function, Oxford University Press, 3 edition, 2003.

[4] J. Mekyska, Z. Smekal, M. Kostalova, M. Mrackova, S. Skutilova, I. Rektorova, Motor aspects of speech imparment in parkinson's disease and their assessment, Cesk Slov Neurol N 74 (2011) 662–668.

[5] A. M. Goberman, Correlation between acoustic speech characteristics and non-speech motor performance in parkinson disease, Med Sci Monit 11 (2005) CR109–116.

[6] A. Tsanas, M. A. Little, P. E. McSharry, L. O. Ramig, Nonlinear speech analysis algorithms mapped to a standard metric achieve clinically useful quantification of average Parkinson's disease symptom severity, J R Soc Interface 8 (2010) 842–855.

[7] P. Boersma, D. Weenink, Praat: Doing Phonetics by Computer (Version 5.3.10), 2012.

[8] A. Tsanas, M. A. Little, P. E. McSharry, A methodology for the analysis of medical data, Chapter 7 in Handbook of Systems and Complexity in Health, Springler, 2013.